Network identification with partial excitation and measurement

Alexandre S. Bazanella, *Senior Member*   Michel Gevers, *Life Fellow*   Julien M. Hendrickx

Abstract—The question addressed in this paper is to determine what is the input/output information that must be acquired in order to identify a network of dynamical systems, or part of such network. More specifically, which nodes must be excited and which nodes must be measured in order to render a network, or a given piece of a network, identifiable. We show that identification of a given set of edges requires that all nodes connected to these edges must be either excited or measured or both. We study in detail the identifiability of four types of edge sets: edges arriving at a given node, edges leaving a given node, edges forming a tree and edges forming a loop. For each case, we characterise excitation and measurement patterns that provide identifiability.

I. INTRODUCTION

This paper deals with the identifiability of dynamical networks in which the node signals are connected by causal linear time-invariant transfer functions and are excited by known external excitation signals. Such networks can be looked upon as connected directed graphs in which the edges between the nodes (or vertices) are scalar transfer functions, and in which known external excitation signals enter into the nodes.

Over the last few years, a significant literature has been developed about the identifiability of such networks [1]–[7]. A number of conditions for the identifiability of the whole network have been derived under prior assumptions on the structure of the network, involving either its external excitation structure, or possibly also its internal structure [5], [6]. In addition to the identification of the whole network, a number of papers have also adressed the identifiability of a single module, or of a subset of modules, typically the incoming edges or the outgoing edges of a given node [6], [8], [9].

Up until recently, the common assumption in this work on identifiability of networks was that all nodes were measured. In [7], [10], this assumption was relaxed and a number of new results were developed for the situation where only a subset of nodes are measured, but all nodes are excited. These results were expanded in [11]. In both these papers it is assumed that the topology of the network is known: one knows which nodes are connected by directed edges, which

Alexandre S. Bazanella is with the Department of Automation and Energy, Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, Brazil, bazanella@ufrgs.br.

Michel Gevers and Julien M. Hendrickx are with the ICTEAM institute, UCLouvain, Louvain la Neuve, Belgium, julien.hendrickx@uclouvain.be, michel.gevers@uclouvain.be.

corresponds to knowing which of the transfer functions of the network matrix describing the connections between the nodes are nonzero. A number of necessary and sufficient conditions for identifiability of a particular edge, or of a set of edges, that leave a specific node have been obtained in [11]. An important contribution of that paper is that these results are described in terms of the existence of mutually vertex disjoint paths between the measured nodes and the out-neighbors of the node of interest. Thus, identifiability can be checked by inspection of the local topology of the network around the node of interest. The necessary and sufficient conditions just described have led to a number of other results in [11] in the form of either necessary conditions or sufficient conditions.

All the results of [11] have been obtained under the assumption that only a subset of nodes are measured but that all nodes are excited by sufficiently rich known external excitation signals, so that the closed-loop transfer matrix from all nodes to the measured nodes can be exactly identified.

In this paper, we relax that assumption, and present identifiability results for situations where not all nodes are measured and not all nodes are excited. In the spirit of [11], these results are all in the form of conditions on the local topology, with the overall goal of trying to minimize the excitation/measurement requirements. They can be subdivided into three parts.

After stating the problem formally in Section II, we first present in Section III easy to check but important necessary conditions for the identifiability of the network : (i) all source nodes must be excited, and all sink nodes must be measured (See Section II for the definition of sources and sinks); (ii) each node must be either excited or measured.

Our next set of results, presented in Section IV, are an extension of results in [11]: they consider the identifiability of some or all of the outgoing edges of a node on the basis of a subset of measured nodes, but without the assumption that all nodes are excited. A sufficient condition is derived first, which is based on the local topology and on knowledge of a subset of the closed-loop transfer functions that relate all nodes to each other. The dual of these results are also derived: they consider the identifiability of the incoming edges of a given node on the basis of a subset of excited nodes, when not all nodes are measured. In passing, our results allow us to show that a single embedded module (i.e. edge) may be identifiable by simply exciting its input and output node, and measuring its output node, provided some topological properties are satisfied.

Our third set of results, presented in Section V, concern the identifiability of parts of the network that have the structure of

a tree or of an isolated loop[1]. We show that a tree is identifiable if and only if all its sources are excited, all its sinks are measured, and every other node is either excited or measured. For isolated loops we present a simple sufficient condition stating that all transfer functions in such loop are identifiable if one node in the loop is both excited and measured, and all others are either excited or measured. In addition, we show that an isolated loop with an even number of nodes, larger than 3, is identifiable if all nodes in the loop are alternately excited or measured. We conclude in Section VI.

II. STATEMENT OF THE PROBLEM

We consider the identification of elements of a network matrix $G(z)$, where the network is made up of n nodes, with node signals denoted $\{w_1(t), \ldots, w_n(t)\}$, and where these node signals are related to each other and to external excitation signals $r_j(t), j = 1, \ldots, n$ by the following network equations, which we call the **network model**:

$$w(t) = G(z)w(t) + Br(t) + v_1(t) \quad (1)$$
$$y(t) = Cw(t) + v_2(t) \quad (2)$$

The matrix B is a binary selection matrix having full column rank, and each of its columns contains one 1 and $n-1$ zeros. The matrix C is a binary selection matrix having full row rank, and each of its rows contains one 1 and $n-1$ zeros. These matrices define which of the n nodes are excited and which are measured, respectively. $y(t)$ is the vector of measured nodes, while $v_1(t)$ and $v_2(t)$ are possible noise vectors. The matrix $G(z)$ is called the **network matrix**. To the network matrix one can associate a directed graph, in which a directed edge (j, i) is present if $G_{ij}(z) \neq 0$. Thus, the graph defines the topology of the network. The network matrix is assumed to have the following properties:

- its diagonal elements are zero.
- its off-diagonal elements $G_{ij}(z)$ are proper rational transfer functions.
- it is well-posed: $(I - G(z))^{-1}$ is proper and stable [12].
- the graph defined by the network matrix is weakly connected[2].

The selection matrices B and C define two subsets $\mathcal{B}$ and $\mathcal{C}$ of nodes corresponding, respectively, to the nodes that are excited by an external signal and those that are measured.

Define
$$T(z) \triangleq (I - G(z))^{-1}. \quad (3)$$

The **input-output model** corresponding to the network model (1) is then given by

$$y(t) = M(z)r(t) + v(t) \quad \text{with} \quad M(z) \triangleq CT(z)B \quad (4)$$

where $v(t)$ is the effect of noises $v_1(t)$ and $v_2(t)$ on the output $y(t)$. To keep things simple, we assume throughout this paper that the vector $r(t)$ is sufficiently rich so that, for all choices of C and B, $M(z)$ can be consistently estimated by standard

open loop MIMO (Mulitple Input Multiple Output) identification techniques from $\{y(t), r(t)\}$ data. Our identifiability results will be based on assuming that, for a given selection of excited and measured nodes, the input-output matrix $M(z)$ has been first consistently estimated from a (possibly infinite) set of sufficiently rich Input-Output (I/O) data. This is always possible, even in the presence of noise.

We now define the generic identifiability of the network matrix $G(z)$ from $\{y(t), r(t)\}$ data (or equivalently from $M(z)$ given the above assumption), and from the knowledge of the graph structure.

Definition 1. *The network matrix $G(z)$ is* generically identifiable *from excitation signals applied to $\mathcal{B}$ and measurements made at $\mathcal{C}$ if, for any rational transfer matrix parametrization $G(P, z)$ consistent with the directed graph associated to $G(z)$, there holds*

$$C(I - G(P,z))^{-1}B = C(I - \tilde{G}(z))^{-1}B \Rightarrow G(P,z) = \tilde{G}(z) \quad (5)$$

for all parameters P except possibly those lying on a zero measure set in $\mathbb{R}^N$, where $\tilde{G}(z)$ is any network matrix consistent with the graph.

The focus of this paper is on the selection of the sets $\mathcal{B}$ and $\mathcal{C}$ that will ensure network identifiability. The notion of generic identifiability, first introduced in [11], is used because certain networks will always be identifiable except if the transfer functions satisfy some very specific equalities; see [11] for a complete discussion and motivation for this notion.

Notations and definitions

Consider the network matrix $G(z)$ of size $n \times n$ defined in (1) and its associated graph. Define the following:

- $\mathcal{W}$ is the set of all n nodes
- $\mathcal{B}$ is the set of m nodes with excitation
- $\mathcal{C}$ is the set of p measured nodes
- $\mathcal{F}$ is the set of sources, i.e. nodes with no incoming edges
- $\mathcal{S}$ is the set of sinks, i.e. nodes with no outgoing edges
- N_i^+, N_i^- are, respectively, the set of out- and in-neighbors of node i, i.e. the set of nodes j for which $G_{ji} \neq 0$ and the set of nodes j for which $G_{ij} \neq 0$.
- $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$.
- For any matrix T, $T_{\mathcal{D},\mathcal{A}}$ denotes the submatrix of T formed by the rows defined by a set $\mathcal{D}$ and the columns defined by a set $\mathcal{A}$.

It follows from these definitions that $m \leq n$ and $p \leq n$. We shall also use the notion of *vertex disjoint paths*: a group of paths in a graph are mutually vertex disjoint if no two paths of this group contain the same vertex (such a group is sometimes referred to as a linking). Consider two subsets of nodes $\mathcal{A}$ and $\mathcal{B}$. We then denote by $b_{\mathcal{A} \to \mathcal{B}}$ the maximum number of mutually vertex disjoint paths starting in $\mathcal{A}$ and ending in $\mathcal{B}$.

III. NECESSARY CONDITIONS FOR IDENTIFIABILITY

We first recall a fact proved in [10].

Theorem III.1. *The network matrix $G(z)$ is generically identifiable from excitation signals applied to $\mathcal{B}$ and measurements made at $\mathcal{C}$ only if $\mathcal{F} \subseteq \mathcal{B}$ and $\mathcal{S} \subseteq \mathcal{C}$, i.e. all sources are excited and all sinks are measured.*

---

[1]A loop whose nodes do not belong to another loop.

[2]A directed graph is weakly connected if, for any partition of its vertices in two sets, there is at least one edge starting in one of the sets and ending in the other one.

We now provide a necessary condition for the identfiability of the in-neighbors and the out-neighbors of a given node.

Theorem III.2. *Suppose a node is neither excited nor measured. Then none of the corresponding ingoing or outgoing edges can be identified.*

Proof. Without loss of generality let the last node, n, be neither excited nor measured. Now partition the vector w of node signals into a subvector w_1 made up of its first $n-1$ elements and the element w_2 which is the last element of w. Partition $G(z)$ correspondingly; it has the form

$$G(z) = \left(\begin{array}{cc} G_{11} & G_{12} \\ G_{21} & 0 \end{array} \right). \tag{6}$$

Correspondingly, partition $T(z)$ as

$$T(z) \triangleq (I - G(z))^{-1} = \left(\begin{array}{cc} T_{11} & T_{12} \\ T_{21} & T_{22} \end{array} \right). \tag{7}$$

Now observe that from the I/O data we identify $y(t) = M(z)r(t)$ exactly, where

$$M(z) = C_1 T_{11}(z) B_1, \tag{8}$$

where C_1 and B_1 are the selection matrices that define the measured nodes and the excited nodes, respectively, among the first $n-1$ nodes. It follows from (6)-(7) that $T_{11}(z)$ takes the form $T_{11} = [I - G_{11} - G_{12}G_{21}]^{-1}$. Now consider a $\bar{G}(z)$ defined as

$$\bar{G}(z) \triangleq \left(\begin{array}{cc} G_{11} & \bar{G}_{12} \\ \bar{G}_{21} & 0 \end{array} \right) \triangleq \left(\begin{array}{cc} G_{11} & G_{12}Q \\ Q^{-1}G_{21} & 0 \end{array} \right) \tag{9}$$

where Q is any real, complex, or rational scalar transfer function $Q(z)$. It then follows that

$$\bar{T}_{11} = [I - \bar{G}_{11} - \bar{G}_{12}\bar{G}_{21}]^{-1} = [I - G_{11} - G_{12}G_{21}]^{-1} = T_{11}$$

Hence G_{12} and G_{21} are not identifiable because they produce the same $M(z)$ as $\bar{G}_{12}$ and $\bar{G}_{21}$ in (8). $\square$

An important consequence of Theorem III.2 is that the network matrix $G(z)$ can be identified only if each node is either excited or measured. This extends directly to generic identifiability, as stated in the next corollary. A similar result was proved for the case of graphs represented by state-space representations with positive off-diagonal elements in [13].

Corollary III.1. *The network matrix $G(z)$ is generically identifiable from excitation signals applied to $\mathcal{B}$ and measurements made at $\mathcal{C}$ only if $\mathcal{B} \neq \emptyset$, $\mathcal{C} \neq \emptyset$ and $\mathcal{B} \cup \mathcal{C} = \mathcal{W}$.*

Proof. Without any measurement or excitation, one cannot estimate any element of $T(z)$; the last condition is an immediate consequence of Theorem III.2. $\square$

IV. IDENTIFIABILITY OF ELEMENTS OF A COLUMN OR ROW OF $G(z)$

In this section we present results on the identifiability of outgoing and, respectively, incoming edges of a given node. This corresponds to elements that are all in the same column, respectively all in the same row, of the network matrix G.

A. Identification of outgoing edges

We first consider the relatively simpler situation where it is desired to identify all outgoing edges of a given node.

Theorem IV.1. *Consider a node i and its out-neighbors N_i^+. If there exists a subset $\mathcal{C} \subset \mathcal{W}$ of measured nodes such that*
  1) $b_{N_i^+ \to \mathcal{C}} = |N_i^+|$,
  2) *the elements of $T_{\mathcal{C}, N_i^+}$ and $T_{\mathcal{C}, i}$ are known from the topology or can be identified from the data,*
then the transfer functions from node i to its out-neighbors N_i^+ can be generically identified from I/O measurements $\{y, r\}$.

Proof. Recall that $b_{N_i^+ \to \mathcal{C}}$ denotes the maximum number of mutually vertex disjoint paths from the set N_i^+ to the set $\mathcal{C}$. Consider the system of equations for transfer matrices $\bar{G}$ consistent with the graph defined by $G(z)$

$$CT(z)(I - \bar{G}(z)) = C, \tag{10}$$

of which $\bar{G}(z) = G(z)$ is an obvious solution since $T(z) = (I - G(z))^{-1}$. We desire to recover $G_{N_i^+, i}$, i.e. the i-th column of G, from CT. In the i-th column of $I - \bar{G}$, the only nonzero elements are the 1 at position (i, i), and the G_{ji} corresponding to the out-neighbors of i. The columns of T corresponding to the zero elements of the i-th column of $\bar{G}$ do not contribute to the computation of $\bar{G}_{ji}$ using (10). The solution set of (10) for the i-th column of $\bar{G}$ can then be rewritten as

$$T_{\mathcal{C}, N_i^+} \bar{G}_{N_i^+, i} = T_{\mathcal{C}, i} - C_i \tag{11}$$

where C_i is the i-th column of the measurement selector matrix C. Note that C_i is zero if node i is not measured. This system always admits at least one solution $\bar{G}_{N_i^+, i} = G_{N_i^+, i}$ since $\bar{G} = G$ is a solution to the initial system (10). The solution $G_{N_i^+, i}$ of this set of equations is unique and can be computed from (11) if and only if $T_{\mathcal{C}, N_i^+}$ has full column rank and if all elements of $T_{\mathcal{C}, N_i^+}$ and $T_{\mathcal{C}, i}$ are known. It has been shown in Proposition V.1 of [11] that $T_{\mathcal{C}, N_i^+}$ has full column rank if and only if condition 1) of the theorem holds. $\square$

We note that condition 2) of Theorem IV.1 is sufficient but not necessary, as will be seen in Example V.1. Also, the theorem extends a result in [11] showing that condition 1) is necessary and sufficient when all nodes in the network are excited, which corresponds to $B = I$ in (4). Condition 1) remains thus necessary in general, when $B \neq I$.

Comment about the meaning of condition 2
Condition 2) of the theorem means that the submatrices $T_{\mathcal{C}, N_i^+}$ and $T_{\mathcal{C}, i}$ can be computed. This can be by identification of the required elements, in which case the condition is a requirement on the presence of the excitation signals necessary for such identification. A simple way - but not the only way - to identify an element T_{ji} is to excite node i and measure node j. However, some columns of $T_{\mathcal{C}, N_i^+}$ may be known a priori on the basis of local information about the graph structure linking the out-neighbors to the measured nodes. In such case, these columns need not be identified and the corresponding nodes need not be excited. The following Lemma can be proved by direct inspection.

Lemma IV.1. *Let $j \in N_i^+$ be an out-neighbor of node i. If there is no loop around node j and if there is no path from this out-neighbor to any measured node in C (except possibly j if it is measured), then $T_{jj} = 1$ and $T_{kj} = 0$ for all $k \in C \setminus j$.*

It follows from this simple result that, for all out-neighbors $j \in N_i^+$ that satisfy the conditions of Lemma IV.1, the corresponding columns of T_{C,N_i^+} are known by the local graph structure and that they therefore need not be identified. This means that the corresponding nodes need not be excited.

A sufficient condition for the identification of all outgoing edges of a given node i is given by the following Corollary.

Corollary IV.1. *Consider a node i with out-neighbors N_i^+. If there exists a subset $C \subset W$ of measured nodes such that*

1) $b_{N_i^+ \to C} = |N_i^+|$
2) $i \in \mathcal{B}$
3) $N_i^+ \subset \mathcal{B}$

then the transfer functions from node i to its out-neighbors N_i^+ can be generically identified from I/O measurements $\{y, r\}$.

Proof. We apply Theorem IV.1. Condition (1) states the existence of at least one set of $|N_i^+|$ vertex disjoint paths from N_i^+ to C. Moreover, conditions 2) and 3) imply that the transfer function matrices T_{C,N_i^+} and $T_{C,i}$ can be consistently estimated from the data. Hence conditions (1) and (2) of Theorem IV.1 are satisfied, and the result follows. $\square$

A simple solution that satisfies the conditions of this theorem consists in measuring all out-neighbors N_i^+, and exciting node i and all its out-neighbors N_i^+. This solution was presented as Theorem 5.1 in [9].

The excitation conditions of Corollary IV.1 are a simple solution for the identification of the outgoing edges of a given node. However, these conditions are not necessary. First, as we have shown in Lemma IV.1, some columns of T_{C,N_i^+} may be known on the basis of local information about the topology, in which case the corresponding nodes need not be excited. But in addition, computation of the elements of T_{C,N_i^+} and $T_{C,i}$ may be achieved from knowledge of transfer functions T_{kl} other than those appearing in T_{C,N_i^+} and $T_{C,i}$, and these other transfer functions are then identified by exciting node signals other than node i and its out-neighbors. We illustrate this with an example.

Example IV.1. Consider the network defined by the following network matrix G:

$$G = \begin{pmatrix} 0 & 0 & 0 & G_{14} & 0 & 0 & 0 \\ G_{21} & 0 & 0 & 0 & 0 & 0 & G_{27} \\ G_{31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & G_{42} & 0 & 0 & 0 & 0 & 0 \\ 0 & G_{52} & G_{53} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{64} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We want to identify the outgoing edges of node 1, i.e. the first column of G. We illustrate Theorem IV.1 and Corollary IV.1 by describing combinations of measured nodes, i.e. C, and excited nodes, i.e. $\mathcal{B}$, that yield identifiability of G_{21} and G_{31}.

(1) Consider first $C = \{2, 3\}$. By Theorem III.2 we know that $\mathcal{B}$ must contain node 1. The system of equations (11) becomes

$$\begin{pmatrix} T_{22} & 0 \\ T_{32} & 1 \end{pmatrix} \begin{pmatrix} G_{21} \\ G_{31} \end{pmatrix} = \begin{pmatrix} T_{21} \\ T_{31} \end{pmatrix}$$

In accordance with Lemma IV.1, the second column of T_{C,N_i^+} is known because there is no loop around node 3 and no path from 3 to 2. Thus G_{21} and G_{31} can be identified with a choice $\mathcal{B} = \{1, 2\}$. Node 3 need not be excited.

(2) Consider now $C = \{3, 6\}$, i.e. out-neighbor 2 is not measured. By Theorem III.2 we know that it must be excited, thus $\mathcal{B}$ must contain at least nodes 1 and 2. Equations (11) yield

$$\begin{pmatrix} T_{32} & 1 \\ T_{62} & 0 \end{pmatrix} \begin{pmatrix} G_{21} \\ G_{31} \end{pmatrix} = \begin{pmatrix} T_{31} \\ T_{61} \end{pmatrix}$$

The second column is again known from structural information, because node 3 has no loop around it and has no path to the other measured node. An excitation scenario $\mathcal{B} = \{1, 2\}$ again yields identifiability.

(3) Next consider $C = \{2, 5\}$. By Theorem III.2 we know that $\mathcal{B}$ must contain nodes 1 and 3. This scenario yields

$$\begin{pmatrix} T_{22} & 0 \\ T_{52} & T_{53} \end{pmatrix} \begin{pmatrix} G_{21} \\ G_{31} \end{pmatrix} = \begin{pmatrix} T_{21} \\ T_{51} \end{pmatrix}$$

The two columns of T_{C,N_i^+} now need to be identified, in addition to $T_{C,i}$. An excitation scenario $\mathcal{B} = \{1, 2, 3\}$ will yield identifiability.

(4) Our final example shows that knowledge of the columns of T_{C,N_i^+} does not necessarily require that the input nodes of the corresponding T_{jk} be excited. Let $C = \{2, 3\}$ as in case (1) but assume that, instead of exciting nodes 1 and 2, we excite nodes 1 and 7, i.e. $\mathcal{B} = \{1, 7\}$. This means that the transfer functions $T_{21}, T_{27}, T_{31}, T_{37}$ can all be consistently estimated. Some straightforward calculations then show that the required T_{22} and T_{32} can be computed from those as follows:

$$T_{22} = \left(1 - \frac{T_{21}T_{37}}{T_{31}T_{27}} \right)^{-1}, \quad T_{32} = \frac{T_{37}T_{22}}{T_{27}}$$

This case shows that all elements of T_{C,N_i^+} can be computed even though condition 3) of Corollary IV.1 is not satisfied. $\blacksquare$

We turn next to the case where one wants to identify only a subset of the outgoing edges of a node, in the situation where not all nodes are excited.

Theorem IV.2. *Consider a node i with out-neighbors N_i^+. Let $N_i^* \subseteq N_i^+$ be a subset of N_i^+ and denote $\bar{N}_i^* \triangleq N_i^+ \setminus N_i^*$. If there exists a subset $C \subset W$ of measured nodes such that*

1) $b_{N_i^* \to C} = |N_i^*|$,
2) $b_{N_i^+ \to C} = |N_i^*| + b_{\bar{N}_i^* \to C}$,
3) *the elements of T_{C,N_i^*}, $T_{C,\bar{N}_i^*}$ and $T_{C,i}$ are known from the topology or can be identified from the data,*

then the transfer functions from node i to N_i^* can generically be identified from I/O measurements.

Proof. Let $G_{N_i^*,i}$ denote the elements of the i-th column of G that correspond to the selected out-neighbors N_i^* and $G_{\bar{N}_i^*,i}$

be the remaining nonzero elements. By splitting $G_{N_i^+,i}$ into these two subvectors, one can now rewrite (11) as

$$T_{\mathcal{C},N_i^*}G_{N_i^*,i} + T_{\mathcal{C},\bar{N}_i^*}G_{\bar{N}_i^*,i} = T_{\mathcal{C},i} - C_i. \quad (12)$$

This set of equations has a unique solution $G_{N_i^*,i}$ if and only if $T_{\mathcal{C},N_i^*}$ has full column rank, the image sets of $T_{\mathcal{C},N_i^*}$ and $T_{\mathcal{C},\bar{N}_i^*}$ do not intersect, and all elements of $T_{\mathcal{C},N_i^*}$, $T_{\mathcal{C},\bar{N}_i^*}$ and $T_{\mathcal{C},i}$ are known. The result then follows from the proof of Theorem V.1 of [11] where it has been shown that the first of these conditions is equivalent to condition 1) of the theorem while the second condition is equivalent to condition 2). $\square$

We note again that conditions 1) and 2) were shown in [11] to be necessary and sufficient for identifiability when all nodes are excited ($B = I$) and remain thus necessary in general. The next corollary is an interesting special case of Theorem IV.2.

Corollary IV.2. *Consider that it is desired to identify a single module G_{ji} in a network. Let N_i^+ be the out-neighbors of node i and denote by $\bar{N}_i^+ = \{N_i^+ \setminus j\}$, i.e. the set made up of the other out-neighbors of node i. Suppose that these nodes have no path to node j. Then G_{ji} can be identified by exciting nodes i and j and measuring node j, as the solution of*

$$G_{ji} = \frac{T_{ji}}{T_{jj}}. \quad (13)$$

Proof. Consider equation (12) with the choice $\mathcal{C} = j$ and $N_i^* = j$. Under the assumptions of the theorem, the row vector $T_{\mathcal{C},\bar{N}_i^*}$ is zero, while $C_i = 0$ since C selects the j-th row of the identity matrix. Hence (12) reduces to (13). $\square$

Example IV.2. We illustrate this corollary using the same example as before. Suppose we only want to identify G_{21}. Since node 3 has no loop around it and no path to node 2, it follows that G_{21} is obtained simply as the solution of $T_{22}G_{21} = T_{21}$. Hence the identification of the single module G_{21} can be achieved by simply exciting nodes 1 and 2 and measuring node 2.

B. Identification of incoming edges

For a given choice of measured nodes determined by $\mathcal{C}$, the identification of the outgoing edges of a node i have been based on solving equation (10) for its i-th column. In this subsection we examine the dual situation where one wants to identify the $|N_i^-|$ incoming edges of a node i. For a given set $\mathcal{B}$ of excited nodes, the identification of these incoming edges are based on solving the equation

$$(I - G(z))T(z)B = B \quad (14)$$

for its i-th row, where B is the excitation node selection matrix corresponding to the set $\mathcal{B}$. In the i-th row of G, the only nonzero elements are the 1 in position (i,i) and the G_{ij} corresponding to the in-neighbors of i. Thus (14) can be rewritten as

$$G_{i,N_i^-}T_{N_i^-,\mathcal{B}} = T_{i,\mathcal{B}} - B_i \quad (15)$$

where we denote by B_i the i-th row of B. Note that $B_i = 0$ unless node i is excited.

The following results are dual to Theorem IV.1, Corollary IV.1 and Theorem IV.2, and are proved in an exactly parallel way.

Theorem IV.3. *Consider a node i and its in-neighbors N_i^-. If there exists a subset $\mathcal{B} \subset \mathcal{W}$ of excited nodes such that*

1) $b_{\mathcal{B} \to N_i^-} = |N_i^-|$
2) *the elements of $T_{N_i^-,\mathcal{B}}$ and $T_{i,\mathcal{B}}$ are known from the topology or can be identified from the data,*

then the transfer functions to node i from its in-neighbors N_i^- can be generically identified from I/O measurements.

Corollary IV.3. *Consider a node i with in-neighbors N_i^-. If there exists a subset $\mathcal{B} \subset \mathcal{W}$ of excited nodes such that*

1) $b_{\mathcal{B} \to N_i^-} = |N_i^-|$
2) $i \in \mathcal{C}$
3) $N_i^- \subseteq \mathcal{C}$

then the transfer functions to node i from its in-neighbors N_i^- can all be generically identified from I/O measurements $\{y, r\}$.

Theorem IV.4. *Consider a node i with in-neighbors N_i^-. Let $N_i^* \subseteq N_i^-$ be a subset of N_i^- and denote $\bar{N}_i^* \triangleq \{N_i^- \setminus N_i^*\}$. The transfer functions to node i from N_i^* can be generically identified from I/O measurements if the following three conditions are satisfied*

1) $b_{\mathcal{B} \to N_i^*} = |N_i^*|$
2) $b_{\mathcal{B} \to N_i^-} = |N_i^*| + b_{\mathcal{B} \to \bar{N}_i^*}$
3) *the elements of $T_{N_i^*,\mathcal{B}}$, $T_{\bar{N}_i^*,\mathcal{B}}$ and $T_{i,\mathcal{B}}$ are known from the topology or can be identified from the data.*

We note that a simple solution that satisfies the conditions of Corollary IV.3 consists in exciting all in-neighbors N_i^-, and measuring node i and all its in-neighbors N_i^-. This solution was presented as Theorem 5.2 in [9].

V. PARTICULAR STRUCTURES WITHIN A NETWORK

Unlike the situation that prevails in the context of generic identifiability, we assume in this section that $G_{ij}(z)$ cannot be zero if the edge (i,j) is present in the graph.

A. Trees

A tree is a weakly connected graph which has no loops even if one were to change the edges directions. Direct consequences of this definition are the following:

1) for any pair of nodes i and j there is at most one path from i to j;
2) if there is a path from node i to node j, then there is no path from j to i;
3) if i is an in-neighbor of j then $T_{ji} = G_{ji}$;
4) if there is a path from node i to node j passing through node k then $T_{ji} = T_{jk}T_{ki}$.

A tree is composed of branches: a branch is a walk $\mathcal{P} = \{k_1, \ldots, k_m\}$ between a source k_1 and a sink k_m. All nodes of a tree belong to at least one branch. We now state a necessary and sufficient condition for the identifiability of a tree.

Theorem V.1. *A tree is identifiable if and only if the following conditions are satisfied*

1) $\mathcal{F} \subseteq \mathcal{B}$: *all sources are excited*
2) $\mathcal{S} \subseteq \mathcal{C}$: *all sinks are measured*
3) $\mathcal{B} \cup \mathcal{C} = \mathcal{W}$: *each node is either measured or excited.*

Proof. Necessity: We know from Corollary III.1 that every node must either be measured or excited, which is condition 3. Moreover, if a source is not excited, then the identified closed loop transfer functions present in $CT(z)B$ are totally independent of the transfer function corresponding to the edge leaving that source, and therefore do not allow recovering it. The same holds true if a sink is not measured.

Sufficiency: Because the sources are excited and the sinks are measured, for each branch $\mathcal{P} = \{k_1, \ldots, k_m\}$ between a source k_1 and a sink k_m, the transfer functions $T_{k_m k_1}$ between them are known and given by

$$T_{k_m k_1} = G_{k_m k_{m-1}} G_{k_{m-1} k_{m-2}} \ldots G_{k_3 k_2} G_{k_2 k_1}$$

In general, for all $j > i$ we have:

$$T_{k_j k_i} = G_{k_j k_{j-1}} \ldots G_{k_{i+1} k_i}.$$

Now take node k_2. If it is measured, then we know $T_{k_2 k_1}$ which is equal to $G_{k_2 k_1}$; if it is excited then we know $T_{k_m k_2}$ and $G_{k_2 k_1}$ can be obtained as

$$G_{k_2 k_1} = \frac{T_{k_m k_1}}{T_{k_m k_2}}.$$

Next, take node k_3. If it is measured, then we know $T_{k_3 k_1}$, and then $G_{k_3 k_2}$ can be obtained as

$$G_{k_3 k_2} = \frac{T_{k_3 k_1}}{G_{k_2 k_1}};$$

if k_3 is excited, then we know $T_{k_m k_3}$, and $G_{k_3 k_2}$ can be obtained as

$$G_{k_3 k_2} = \frac{T_{k_m k_1}}{G_{k_2 k_1} T_{k_m k_3}}.$$

Applying this same reasoning sequentially for all nodes in the branch, for all branches in the tree, proves the theorem. $\square$

The same properties hold for a tree that is a sub-graph of a larger graph, provided that all paths leaving the tree never come back to it (and vice-versa).

B. Loops

A "loop graph" is a graph that consists of a single loop and nothing more. Its network matrix is in the form

$$G = \begin{bmatrix} 0 & 0 & \ldots & 0 & G_{1n} \\ G_{21} & 0 & \ldots & 0 & 0 \\ 0 & G_{32} & \ldots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \ldots & G_{n(n-1)} & 0 \end{bmatrix} \quad (16)$$

The loop we are interested in can be a graph by itself, as in (16), or part of a larger graph. When the loop of interest is part of a larger graph, some of its nodes may belong to other loops. When this is not the case - that is, no other loop in the graph contains any of the nodes of the loop of interest - we will say that it is an *isolated loop*. All results in this subsection pertain to the identifiability of isolated loops.

The necessary condition $\mathcal{B} \cup \mathcal{C} = \mathcal{W}$ means that all nodes must be involved in the identification process: they must be either measured or excited. The question is then whether we can identify the loop under this minimum excitation/measurement condition - as we can for trees - or if we need to have nodes that are both measured and excited. Our main result about identifiability of loops states that a loop graph is identifiable if at least one node is both excited and measured.

Theorem V.2. *All transfer functions in an isolated loop are identifiable if $\mathcal{B} \cup \mathcal{C} = \mathcal{W}$ and $\mathcal{B} \cap \mathcal{C} \neq \emptyset$.*

Proof. Assume, without loss of generality, that the node indices go from 1 to n and that the arrows go from i to $i+1$ in the cycle, as in (16). Let $P = G_{1n} G_{n,n-1} \ldots G_{32} G_{21}$ be the product of all transfer functions in the loop. Observe that the closed loop transfer function from one node to itself is

$$T_{ii} = R := (1 - P)^{-1} \quad (17)$$

For distinct i, k, we also let

$$P_{ik} = G_{i,i-1} G_{i-1,i-2} \ldots G_{k+1,k} \quad \text{if} \quad k < i$$
$$P_{ik} = G_{i,i-1} G_{i-1,i-2} \ldots G_{1n} G_{n,n-1} \ldots G_{k+1,k} \quad \text{if } k > i.$$

Observe that the closed-loop transfer function from k to i is

$$T_{ik} = P_{ik} R \quad (18)$$

Moreover, by definition, for every distinct k and i, there holds

$$P = P_{ki} P_{ik}. \quad (19)$$

We let a be a node that is both measured and excited. We therefore know $T_{aa} = R = (1 - P)^{-1}$ by (17), and hence P. Consider now an arbitrary node i. If it is measured, we know $T_{ia} = P_{ia} R$ by (18), and hence P_{ia}, since we know R. The knowledge of P together with (19) gives us P_{ai}. Similarly, if i is excited, we know T_{ai}, and we can recover P_{ai} and P_{ia}. We can then sequentially recover P_{ai} and P_{ia} for all i and then all the $G_{i+1,i}$ using the relation $P_{i+1,a} = G_{i+1,i} P_{ia}$. $\square$

Notice from the proof that only one node (any node) needs to be both measured and excited, and how this node is of fundamental importance, for it provides the loop transfer function P. The next theorem shows that, when the number of nodes in the loop is even, there may be an alternative way to determine P, which corresponds to the minimum excitation/measurement condition mentioned above. Thus, the sufficient condition $\mathcal{B} \cap \mathcal{C} \neq \emptyset$ of Theorem V.2 is not necessary.

Theorem V.3. *Let n be even and larger than 3. All transfer functions in an isolated loop can be identified if its nodes are alternately measured and excited.*

Proof. We assume without loss of generality that nodes are numbered sequentially in the loop, with odd nodes being excited and even ones measured. Then the closed-loop transfer functions that are identified are all T_{ji} with j even and i odd. For each even node k, the closed-loop transfer functions from its in-neighbor $k - 1$ and out-neighbor $k + 1$ are

$$T_{k,k-1} = G_{k,k-1} R \quad (20)$$
$$T_{k,k+1} = \frac{P}{G_{k+1,k}} R. \quad (21)$$

We now form the product

$$\prod_{k \text{ even}} \frac{T_{k,k+1}}{T_{k,k-1}} = \prod_{k \text{ even}} \frac{P}{G_{k+1,k}G_{k,k-1}}$$

$$= \frac{P^{\frac{n}{2}}}{\prod_{k \text{ even}} G_{k+1,k}G_{k,k-1}} = \frac{P^{\frac{n}{2}}}{P} \quad (22)$$

and thus P can be calculated from the identified transfer functions. Once P is known, each G_{ji} arriving at an even node - that is, the $G_{k,k-1}$ - can be calculated from (20) and each G_{ji} leaving an even node - that is, the $G_{k+1,k}$ - can be calculated from (21). $\qquad\square$

The following example shows that Theorem V.3 does not hold without the interleaving condition. It also illustrates that condition 2) of Theorem IV.1 is not a necessary condition.

Example V.1. Consider a 4-node cycle where we excite nodes 1, 2 and measure nodes 3, 4. Denoting again $P = G_{21}G_{32}G_{43}G_{14}$ and $R = (1 - P)^{-1}$, we now identify

$$T_{31} = G_{32}G_{21}R \quad (23)$$
$$T_{32} = G_{32}R \quad (24)$$
$$T_{41} = G_{43}G_{32}G_{21}R \quad (25)$$
$$T_{42} = G_{43}G_{32}R \quad (26)$$

Observe that G_{32} and R always appear together, and G_{14} only appears as part of R; define $G_{32}R =: Q$. Any modification of G_{32} to G'_{32} and G_{14} to G'_{14} that leaves $G'_{32}R' = Q$ (for R' defined analogously to R) is undetectable. Observe now that $G'_{32}R' = Q$ is equivalent to

$$G'_{32} = Q(1 - P') = Q - QG_{21}G'_{32}G_{43}G'_{14}$$

or

$$G'_{32} = \frac{Q}{1 - QG_{21}G_{43}G'_{14}}$$

Hence for any arbitrary G'_{14}, we can find a G'_{32} that leaves $G'_{32}R' = Q$ and leads thus to the same identified closed loop transfer function as our initial transfer functions G_{14}, G_{32}. The network is therefore not identifiable, and we can in particular not recover G_{14} and G_{32}. $\qquad\blacksquare$

This example also shows that the sufficient condition in Theorem IV.1 is not necessary. Dividing (26) by (24) shows that $G_{43} = T_{42}/T_{32}$ can be recovered. However, condition (2) of Theorem IV.1 does not hold for $\mathcal{C} = \{3, 4\}$. Indeed we argue that we cannot recover T_{33} nor T_{43} from the chosen excitation/measurement pattern, even though we actually know the ratio between these quantities. Observe that $T_{33} = 1 + P + P^2 + \ldots = (1 - P)^{-1} = R$. So if we knew T_{33} we would know R, from which we could deduce G_{32} via (24). But we have proved above that G_{32} cannot be recovered. Similarly, one can verify that $T_{43} = RG_{43}$, so the knowledge of T_{43} together with our ability to recover G_{43} would again allow deducing R, which we have argued is impossible.

Notice also how n being even is crucial for the validity of Theorem V.3. If n is odd, and under the remaining conditions

of the Theorem, then nodes 1 and n are neighbors that are both excited and not measured. This violation of the interleaving condition implies that we cannot find P (thus R) in the same way as in the case of n even, using (22), because now the last G is missing in the product: $\prod_{k \text{ even}} G_{k+1,k}G_{k,k-1} = (P/G_{1,n}) \neq P$.

VI. Conclusions

In [11] we had analyzed in detail the problem of which nodes must be measured to identify parts of a network where all the nodes are excited with persistently excited signals. In this paper we have extended in some directions the results of [11] to the case where not all nodes are excited. We have given conditions on the excitation and measurement pattern that provide identifiability for different sets of edges: a row of the network matrix, a column of the network matrix, a tree and a loop. For trees we have provided a necessary and sufficient condition, whereas for loops we have given a tight sufficient condition - which requires only one additional excitation or measurement with respect to the minimal excitation/measurement requirement imposed by the necessary condition. For the case of a column - and its dual of a row - we have also given tight sufficient conditions, but a general characterisation of the necessary and sufficient conditions in these cases has so far shown to be elusive.

References

[1] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1860–1870, 2010.

[2] A. Dankers, P. Van den Hof, P. Heuberger, and X. Bombois, "Dynamic network structure identification with prediction error methods - basic examples," in *USB Proc. 16-th IFAC Symp on System Identification (SYSID 2012)*. IFAC, 2012, pp. 876–881.

[3] A. Chiuso and G. Pillonetto, "A Bayesian approach to sparse dynamic network identification," *Automatica*, vol. 48, pp. 1553–1565, 2012.

[4] H. Weerts, A. Dankers, and P. Van den Hof, "Identifiability in dynamic network identification," in *USB Proc. 17th IFAC Symp. on System Identification*, Beijing, P.R. China, 2015, pp. 1409–1414.

[5] M. Gevers, A. Bazanella, and A. Parraga, "On the identifiability of dynamical networks," in *USB Proc. IFAC World Congress 2017*. Toulouse, France: IFAC, July 2017, pp. 11 069–11 074.

[6] H. Weerts, P. Van den Hof, and A. Dankers, "Identifiability of linear dynamic networks," *Automatica*, vol. 89, pp. 247–258, 2018.

[7] H. van Waarde, P. Tesi, and M. Camlibel, "Topological conditions for identifiability of dynamical networks with partial node measurements," in *IFAC PapersOnLine*, IFAC, Ed., vol. 51, no. 23, 2018, pp. 319–324.

[8] H. Weerts, P. Van den Hof, and A. Dankers, "Single module identifiability in linear dynamic networks," in *Accepted for presentation at 57th IEEE Conference on Decision and Control (CDC2018)*, IEEE, Ed., Miami Beach, Florida, USA, December 2018.

[9] A. Bazanella, M. Gevers, and G. V. da Silva, "A practical method for the consistent identification of a module in a dynamical network," in *USB Proc. IFAC Symposium on System Identification (SYSID 2018)*. Stockholm, Sweden: IFAC, July 2018, pp. 862–867.

[10] A. S. Bazanella, M. Gevers, J. M. Hendrickx, and A. Parraga, "Identifiability of dynamical networks: Which nodes need be measured?" in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 5870–5875.

[11] J. Hendrickx, M. Gevers, and A. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Transactions on Automatic Control; doi:10.1109/TAC.2018.2867336*, 2018.

[12] M. Araki and M. Saeki, "A quantitative condition for the well-posedness of interconnected dynamical systems," *IEEE*, vol. 28, no. 5, pp. 625–637, May 1983.

[13] H. van Waarde, P. Tesi, and M. Camlibel, "Identifiability of undirected dynamical networks: a graph-theoretic approach," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 683–688, 2018.